Multi-center validation of an automatic algorithm for fast 4D myocardial segmentation in cine CMR datasets

Sandro Queirós[a,b,c,*] , Daniel Barbosa[b,d], Jan Engvall[e], Tino Ebbers[e], Eike Nagel[f], Sebastian I. Sarvari[g], Piet Claus[a], Jaime C. Fonseca[c], João L. Vilaça[b,d], and Jan D'hooge[a]

[a] Lab on Cardiovascular Imaging and Dynamics, KU Leuven, Belgium
[b] ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal
[c] Algoritmi Center, School of Engineering, University of Minho, Guimarães, Portugal
[d] DIGARC – Polytechnic Institute of Cávado and Ave, Barcelos, Portugal
[e] Department of Clinical Physiology, Department of Medical and Health Sciences, Department of Science and Technology, Center for Medical Image Science and Visualization (CMIV), Linköping University, Linköping, Sweden
[f] Cardiovascular Imaging Department, Division of Imaging Sciences, King's College London, London, United Kingdom
[g] Department of Cardiology and Center for Cardiological Innovation, Oslo University Hospital, Rikshospitalet, Oslo, Norway

*Corresponding author:
Sandro Queirós
E-mail address: sandroqueiros at ecsaude.uminho.pt
Address: Life and Health Sciences Research Institute (ICVS)
Universidade do Minho
Campus de Gualtar
4710-057 Braga
Portugal

Abstract

Aims: Quantitative analysis of cine cardiac magnetic resonance (CMR) images for the assessment of global left ventricular (LV) morphology and function remains a routine task in clinical cardiology practice. To date, this process requires user interaction and therefore prolongs the examination (i.e. cost) and introduces observer variability. In this study, we sought to validate the feasibility, accuracy and time-efficiency of a novel framework for automatic quantification of LV global function in a clinical setting.

Methods and Results: Analyses of 318 CMR studies, acquired at the enrollment of patients in a multi-center imaging trial (DOPPLER-CIP), were performed automatically, as well as manually. For comparative purposes, intra- and interobserver variability was also assessed in a subset of patients. The extracted morphological and functional parameters were compared between both analyses, and time-efficiency was evaluated. The automatic analysis was feasible in 95% of the cases (302/318) and showed a good agreement with manually-derived reference measurements, with small biases and narrow limits of agreement particularly for end-diastolic volume (-4.08 ± 8.98 ml), end-systolic volume (1.18 ± 9.74 ml) and ejection fraction (-1.53 ± 4.93 %). These results were comparable to the agreement between two independent observers. A complete automatic analysis took 5.61 ± 1.22 seconds, which is nearly 150 times faster than manual contouring (14 ± 2 min, $p \leq 0.05$).

Conclusion: The proposed automatic framework provides a fast, robust and accurate quantification of relevant LV clinical indices in "real-world" cine CMR images.

Keywords: cardiac cine MRI; left ventricular quantification; clinical validation; automatic segmentation; fast image processing.

Introduction

The assessment of left ventricular morphology and function using non-invasive imaging is routinely performed for the diagnosis and treatment follow-up of patients with cardiovascular diseases. Among the available imaging techniques, cardiac magnetic resonance (CMR) imaging is a successful and promising modality with excellent accuracy and reproducibility for cardiac morphologic and functional assessment [1].

To extract clinically relevant information, accurate post-processing of the cine CMR images is required. In clinical routine, manual contouring is usually performed for the quantification of left ventricular volumes, mass (LVM) and ejection fraction (LVEF). To this end, the physician must delineate the left ventricle (LV) by contouring the endo- and epicardial boundaries in the slices covering the LV, i.e. from the atrioventricular ring to the apex, at both the end-diastolic (ED) and end-systolic (ES) cardiac phases [2]. Such task is tedious, time consuming and unpractical [3], normally requiring between 6 and 20 min for a complete analysis [2-4]. In addition, manual contouring is prone to intra- and interobserver variability, mainly associated with the choice of ED and ES phases; the choice of the most basal LV slice; and the chosen endocardial edge selection approach [2, 5]. In fact, according to the recent study of Miller *et al.* [2], the analysis methodology used is a major source of variability in a "real-world" scenario, leading to significant differences in the quantified LV indices. In practice, specific reference ranges would have to be established depending on the contouring methodology employed by the physician, which is cumbersome and impractical.

To deal with these issues and ultimately obtain more confident, accurate and robust quantifications, automatic delineation of the myocardium in cine CMR images has been pursued over the last decades. To date, several methods have been presented for (semi-)automated LV CMR segmentation [3]. In addition, several commercial software packages are currently available for manual, semi-automated or automated analysis of these datasets [5-11]. Notwithstanding, due to the difficulties in designing a solution able to deal with the segmentation challenges present in CMR images [12], fast, automatic and accurate assessment of LV boundaries from base to apex is still lacking, often requiring manual correction of imperfect contours [3, 4, 13]. Thus, automatic LV CMR segmentation remains an open problem with significant clinical value,

supported by the fact that it receives significant amount of attention in contemporary literature [12, 14-18]. Besides the accuracy and robustness of a given segmentation methodology, the time required to segment is also a critical factor for success when introducing a given approach in daily clinical practice.

In this study, we sought to assess the feasibility, accuracy and time-efficiency of a fast automatic LV segmentation framework, originally presented in (12), in a multi-center database of CMR images mimicking, as close as possible, "real-world" clinical variability in image quality and anatomical and functional characteristics.

Methods

Study population

The study population consisted of 318 patients (236 men and 82 women; age: 64 ± 8.3 years; range: 35–83 years) with suspicion of ongoing (chronic) myocardial ischemia. These patients were enrolled in a multi-center imaging trial (Determining Optimal non-invasive Parameters for the Prediction of Left vEntricular morphological and functional Remodeling in Chronic Ischemic Patients, DOPPLER-CIP), whose inclusion and exclusion criteria details can be found in (19). Among the DOPPLER-CIP population, these patients represent the subset which underwent CMR image acquisition, and whose core-lab measurements were available at the time of this study. Table 1 details the clinical characteristics of the overall population at the time of acquisition. All patients gave written informed consent prior to CMR imaging. The study was carried out in accordance with the guidelines of the ethics committee of each of the sites involved.

Image acquisition

Cardiac MR images were acquired at 4 different sites using different acquisition protocols and MR imaging systems, following recommendations of the Society for Cardiovascular Magnetic Resonance [20]. All studies were performed with subjects in the supine position. A stack of short-axis (SAX) images, covering the LV from base to apex, was acquired using an ECG-triggered balanced steady state free-precession (b-SSFP) pulse sequence, during consecutive end-expiratory breath-holds. Typical CMR

parameters for the different sites can be found in Table 2. These settings were optimized for every patient, resulting in 181 unique acquisition profiles and, thus, reflecting the variation in image settings that can be found in clinical practice.

Image analysis

As a preliminary step, in order to define the same set of images to be manually and automatically analyzed, the stack of SAX slices was cropped to only include the slices effectively covering the LV, from the atrioventricular ring to the apex. The most basal slice was visually defined as the slice having at least fifty percent of the LV cavity surrounded by myocardium. The apical slice was defined as the most apical slice still showing an intra-cavitary blood pool [(21)]. Both basal and apical slices were independently defined at the ED and ES phases. In turn, ED and ES phases were visually identified as the frames with maximal and minimal ventricular cross-sectional areas at the mid-ventricular level, respectively [(21)]. For both manual and automatic analyses, trabeculae and papillary muscles (TPMs) were considered part of the LV volume [(22)].

For each methodology, the end-diastolic volume (EDV) and end-systolic volume (ESV) were calculated according to the Simpson's method [(2)], *i.e.* summing up the area enclosed by the endocardium multiplied by the inter-slice distance, in all the slices imaged for ED and ES, respectively. From these two volumetric parameters, stroke volume (SV) and LVEF were calculated, as presented in equations (1) and (2). Finally, LV mass (LVM) was obtained by multiplying the myocardial volume (volume between endocardial and epicardial boundaries, V_{myo}) with the specific density of the myocardial tissue (equation (3)) [(23)].

$$SV(ml) = EDV - ESV \quad (1)$$

$$LVEF(\%) = (EDV - ESV) \times 100/EDV \quad (2)$$

$$LVM(g) = V_{myo} \times 1.05 g/cm^3. \quad (3)$$

Manual analysis

All datasets were analyzed in a core-lab facility whose expertise focuses on MR imaging (Linköping University, Sweden[(19)]). For each patient's dataset, both endo and epicardial contours were manually drawn

at ED and ES, in the SAX images defined above, using the free research segmentation software Segment version 1.9 R2966 (http://www.medviso.com) [24]. Note that the left ventricular outflow tract (LVOT) was excluded from the blood pool during contouring. All manual delineations, as well as the pre-processing slices/frames selection step, were performed by an imaging technician with several years of experience, blinded to the automatic analysis and under the supervision of a level 3 certified (EACVI) MR cardiologist with more than 15 years of experience. The latter was also responsible for the assessment of the image quality according to recommendations [25].

To establish the reproducibility of the manual analysis, a second observer (level 3 certified MR cardiologist) re-analyzed 32 datasets (10% of the database). Moreover, 15 datasets were re-analyzed, with a 1 year interval, in order to assess intraobserver variability as well. In both intra- and interobserver variability assessments, the datasets were randomly selected as a subset of each site, keeping their relative distribution in the database (see Table 2). Note that all analyses were performed on the cropped data, thus having the same implicit variability as when compared to the automatic analysis.

Automated analysis

Automatic ventricular analysis was performed using our recently proposed LV segmentation framework [12]. In brief, the framework delineates the myocardial boundaries throughout the cardiac cycle in a three-step approach: (1) detection of the LV myocardium in an end-diastolic mid-ventricular SAX slice; (2) 3D myocardial segmentation at the ED phase; and (3) tracking of the ED contours throughout the cardiac cycle.

The first block uses the specific appearance of the myocardium in cine CMR data to automatically detect the blood pool and the myocardial annular shape. After detection, an automatic border detection algorithm is employed, whose key principle is to use local contrast and edge information as attractors of the LV surface. In other words, after (automatic) initialization of the LV boundary position, the 3D surface is automatically deformed towards positions of high contrast, indicating blood-tissue or tissue-outside interfaces. While doing so, the framework guarantees the smoothness and the spatial coherence of the 3D

surface, allowing to globally optimize the entire surface without introducing local spatial irregularities on the segmented LV shape. In the last block, the LV surface segmented at the ED phase is tracked over all cardiac phases, ultimately obtaining both endo and epicardial boundaries in the ES phase. Hereto, a global anatomical affine optical flow algorithm is used, which allows performing in-plane tracking of the myocardial motion between adjacent frames. Further technical details regarding the proposed LV segmentation framework can be found in (12).

The analysis dataflow is shown in Figure 1. After loading the cropped data on a computer (Matlab code running on an Intel (R) Core (TM) i7 CPU at 3.6 GHz with 16 GB shared memory), no further user input was requested and, once automatically segmented, no manual corrections were applied. From the segmentation result, the relevant LV indices were computed and exported.

Statistical analysis

To quantify the applicability of the proposed framework for automatic LV segmentation, feasibility was assessed by considering a successful analysis when the detected and delineated object roughly corresponded to the LV. In other words, if any of the intermediate steps of the automatic framework failed upon visual assessment (e.g., the LV myocardium detection step identified the right ventricle instead of the LV, or the 3D initialization step failed to properly identify the LV cavity in all slices from base to apex due to substantial inter-slice misalignment), the analysis was considered incorrect. Note that the remaining results only consider the successfully analyzed datasets.

The accuracy of automatically detected myocardial contours was first assessed by measuring the average perpendicular distance (APD) between the automatically segmented contours and the manually defined ones [(12)]. For each dataset, APD was computed slice by slice and an average value for all ED and ES slices were calculated (independently for ED and ES frames, and all together). To establish comparative values, the same analysis was performed between manually delineated contours of the re-analyzed datasets for both the intra- and interobserver variability analyses. These results are summarized as mean ± standard deviation (SD).

In addition, to assess the clinical value of the proposed framework, the automatically computed cardiac indices were compared against the ones obtained through manual contouring. For such agreement assessment, Bland-Altman analysis [26] were performed for end-diastolic, end-systolic and stroke volumes, myocardial mass and ejection fraction, computing both bias (mean difference) and limits of agreement (i.e. 1.96 times the standard deviation). Moreover, differences between automatically and manually computed indices were tested using a two-tailed paired t-test. The same analysis was applied to the re-analyzed datasets to assess intra- and interobserver variability. To compare the limits of agreement of the observer variability analyses against the automatic framework, a two-tailed F-test was used. All analyses considered a p-value ≤ 0.05 statistically significant.

Finally, the time-efficiency of the proposed algorithm was evaluated by measuring the time spent to automatically segment a full 4D dataset (all slices from atrioventricular ring to the apex and all frames), and compared to the manual analysis time. In this sense, the time taken to perform the manual tracings (all slices at both ED and ES moments) was also recorded for a subset of datasets (the ones used for intra- and interobserver analyses). Note that the time spent to import the data and to compute the relevant cardiac indices were not included, since these are equally required in both manual and automatic analyses. Moreover, the initial time to identify both ED and ES moments and to select the correct basal and apical slices are not included either for the same reason. These results are summarized as mean ± SD.

Results

Applicability study

Based on the abovementioned visual criterion, the automated analysis was found to be feasible in 95% of the cases (302/318). From the 16 datasets considered incorrectly analyzed, 12 were related to the LV myocardium detection (3.8%, step 1) and 4 to the intermediate 3D initialization step employed during the 3D segmentation module (1.2%, step 2).

Image quality assessment

According to Klinke *et al.* [25], 14 out of the 318 studies were penalized because of a reduction in cine image quality with an average of 1.8 points. Most frequently, the penalties were due to the absence of the most basal slice or due to shimming artifacts.

Comparison with manual analysis

Figure 2 illustrates the segmentation results in different slices for a representative set of cases covering the range between the best and worst segmentation results. The segmentation results at the ES phase for the same cases are illustrated in Figure 3. Table 3 summarizes the performance of the proposed framework against manual contouring, considering all 302 correctly analyzed datasets. Moreover, APD values for both intra- and interobserver analyses are also presented.

The Bland-Altman analysis of the clinical parameters is presented in Figure 4. An excellent agreement between automatic and manual measurements was found, with narrow limits of agreement particularly for EDV, ESV and LVEF. All clinical parameters presented statistically significant biases, with an underestimation of end-diastolic volume of about 3% and a slight overestimation of end-systolic volume (around 2%), which resulted in a small underestimation of the ejection fraction when compared to manually derived values (approximately 3%). Finally, myocardial mass presented the highest overestimation (about 20%). Figure 5 shows the biases and limits of agreement for the intra- and interobserver variability analyses. As expected, narrower limits of agreement were obtained for the intraobserver analysis when compared to the interobserver one. All these results are summarized in Table 4 (absolute and relative values are reported).

Analysis time

The computational time required to automatically segment a 4D dataset was 5.61 ± 1.22s (approximately 0.03 s/slice). Among the three steps of the proposed framework, the 3D segmentation module had the highest computational burden (corresponding to near 60% of the analysis time), followed by the tracking procedure (around 30%) and finally the initialization step (with near 10% of the total computational time). In its turn, manual contouring took 14 ± 2 min. Note that the preliminary cropping of

each dataset (not included in the automatic or manual analyses time) took, on average, one minute, being near 30s to assess apical and basal slices and another 30s to define the systolic and diastolic time frames.

Discussion

In this study, we sought to demonstrate the robustness, accuracy and time-efficiency of a previously proposed framework for automatic LV segmentation in a clinical setting mimicking routine clinical practice. Understandably, its feasibility, and thus its ability to analyze a given dataset independent of the acquisition system, image quality or subject-specific anatomy and function, is one of the most crucial characteristics to be evaluated to understand the potential added value of the solution. With these aspects in mind, the current validation study was performed in an extensive database of 318 patients. Importantly, the data was acquired at 4 different sites across Europe, with multiple vendors and distinct image settings (see Table 2). In fact, being acquired in a clinical setting, the database presented 181 unique acquisition profiles, which closely reflects the possible variation in image settings found in clinical practice. Across the entire database, image quality was found to be excellent. Note, however, that for the purpose of this study no data selection was made based on image quality. Taking the amount of analyzed data and the significant acquisition variability into account, this study presents an unprecedented validation of an automatic software solution for LV analysis in cine CMR images (see Table 5).

In this large multi-center database, the proposed framework was found to be feasible in 95% of the cases, which is within the range of other frameworks proposed in literature [14, 16, 18], as well as available commercial software packages [10, 27, 28]. Notwithstanding, the present study presents a larger image database, covering a wider range of image acquisition profiles, anatomical variability and with multiple vendor machines. Specifically, the main causes for incorrect analysis were associated with the initialization steps of the framework, and related to sub-optimal image quality, including high heterogeneity inside the LV blood pool and low contrast with the myocardium, or due to substantial inter-slice misalignment.

Among the 302 successfully analyzed datasets, the framework showed accurate results (see Table 3), demonstrating its robustness on an extensive database as illustrated in the examples of Figure 2 and

Figure 3. Note that even for the cases with worst results (Figure 2E), the segmentation at ED presents a good agreement with manual contours. At ES (Figure 3), the biggest differences were found for the epicardial boundary at the basal slices, mainly due to the difficulty in tracking the blood pool motion close to the LVOT (in cases of non-circumferential myocardium) using only our in-slice global affine tracking strategy, but also due to the predominantly longitudinal motion at these SAX slices. When compared to the APD values found between the same observer or two independent observers (Table 3), the automatic framework showed slightly larger APD values for the endocardial boundary, and higher errors at the epicardium (as abovementioned, mostly due to the automatically delineated contours at the ES phase).

Accurate assessment of quantitative LV parameters is crucial for patient management, disease diagnosis, risk stratification or therapy selection in clinical practice [29, 30]. In fact, LVEF is normally considered one of the most meaningful measures of LV global function [3], with clinical value in heart failure, valvular disease and other cardiovascular diseases [31, 32]. Obviously, the accuracy by which EDV and ESV are determined limit the reliability of the deduced parameters, which in turn can have an influence on patient management. In its turn, LVM was proven to have a role in predicting morbidity and mortality in coronary heart diseases [29]. For these reasons, the accuracy of the automatically computed cardiac indices, compared to manually-derived ones, is critically important for translation of any new methodology into daily clinical practice. In the current study, an excellent agreement between automatic and manually-derived LV parameters was obtained (

	APD_{endo} (mm)	**APD_{epi} (mm)**
Automatic vs Manual		
ED	1.52 ± 0.76	2.18 ± 0.91
ES	2.37 ± 0.90	3.42 ± 1.11
All	1.92 ± 0.71	2.76 ± 0.90
Intraobserver analysis		
ED	0.78 ± 0.23	0.85 ± 0.30
ES	0.96 ± 0.20	1.04 ± 0.27
All	0.86 ± 0.19	0.94 ± 0.25
Interobserver analysis		
ED	1.22 ± 0.45	1.25 ± 0.30
ES	1.70 ± 0.64	1.55 ± 0.45
All	1.44 ± 0.41	1.39 ± 0.31

Table 4 and Figure 4). Although statistical significant biases were found, these remained below 3% of under or over-estimation. At the same time, and even though each parameter spans a large range of functional values in this database, narrow limits of agreement (LOA) were obtained, particularly for EDV (-4.08 ± 8.98 ml), ESV (1.18 ± 9.74 ml) and LVEF (-1.53 ± 4.93 %). Myocardial mass showed a lower agreement and a higher bias (20.41 ± 20.38 g). Nevertheless, it is important to notice that, even between two observers, myocardial mass is usually the least reproducible and most variable parameter (see Table 4) [29, 33]. Indeed, Steen *et al.* [29] stated that a poor border detection between lung and abdominal tissue and anterior, lateral and inferior cardiac walls leads to higher variability in border definition between observers. Likewise, many other frameworks in literature report similar results regarding mass computation (Table 5).

When compared to the intraobserver reproducibility (Table 4 and Figure 5), the reported biases and LOAs for the automated measurements were found to be larger. In what concerns interobserver variability, the limits of agreement for EDV, SV and LVM were found to be statistically similar ($p > 0.05$). Regarding ESV and LVEF, slightly larger LOAs were obtained for the automated measurements, although with smaller biases when compared to the interobserver values. As abovementioned, the larger variability is mainly related to the tracking methodology used, but also related to the way the LVOT is handled by the automatic framework. While manual contouring excluded LVOT from the LV blood pool, the proposed segmentation pipeline does not directly addresses the slices having a myocardium appearing as an incomplete annular shape (instead delineating complete elliptical annular shapes). In most cases, errors are partially avoided due to the intra- and inter-slice spatial continuity of the LV surface extracted with the proposed framework. Nevertheless, it is one of the main contributors to the reported APD (Table 3), as well as miss-agreement with manually-derived LV parameters (Table 4, particularly LVM and ESV, which affects the LVEF). Importantly, one should note that, if needed, the proposed framework allows for easy manual editing and interaction after automatic segmentation [34]. Although the limits of agreements for the automated measurements are slightly larger than those found on literature for studies of normal subjects [8, 23] or patients with a specific pathology [9], these are within the range of reported values for studies with both patients and normal subjects [35, 36] or when considering both experienced and un-experienced observers [11, 29, 33, 36-38]. In

the latter case, some studies report that an improved consensus is achieved after setting a common protocol among observers. Together, these observations are in accordance to the larger interobserver variability encountered in "real-world" scenarios as reported by Miller *et al.* [2]. Overall, the current strategy, being an automatic framework without user input, mitigates the problem of intra- or interobserver variability during boundary delineation. Moreover, it has the potential to reduce the dependency of the clinical indices measurements on the observers' experience.

The abovementioned results are similar to the ones reported by other semi-automated, automatic and fully automatic software solutions [7, 10, 14-18, 27, 39, 40] (Table 5). Nonetheless, none of these works used a database with such a large variability in image acquisition. The fact that the proposed algorithm obtained competitive results under more challenging (i.e. more clinically representative) conditions, emphasizes the relevance of the proposed framework. Although the studies of Codella *et al.* [14] and Tufvesson *et al.* [17] are automatic strategies and present narrower limits of agreement for all indices, particularly LVEF, these studies allowed manual interaction after analysis. In the current study, no user interactions were allowed prior or after the automatic segmentation.

Despite the fact that the proposed framework still requires the manual identification of both ED and ES cardiac moments, as well as basal and apical slices (which take near 1 min), the following automated delineation takes less than 6 seconds. This holds an extremely efficient solution, being nearly 150 times faster than manual contouring (5.61 ± 1.22s vs 14 ± 2 min, $p \leq 0.05$). Moreover, the reported analysis time includes the complete LV delineation along the full cardiac sequence, and not ED and ES only, further showing the time-efficiency of the proposed methodology when compared to manual analysis. The associated time saving could thus further improve the cost-effectiveness of CMR imaging in patient management [41]. Since it is not restricted to the two main cardiac moments, it might even allow broader analyses to be performed, thus having the potential to help during disease diagnosis. Together with the abovementioned feasibility and accuracy results, it proves the suitability of the proposed framework for LV global function quantification in daily clinical practice.

In conclusion, this study presents the evaluation of a novel software solution for fast automatic quantification of LV morphology and global function in a large multi-center database with "real-world" clinical variability and subject to blinded core-lab analysis. The proposed framework was found to be feasible in 95% of the cases, showing accurate segmentation results. Importantly, the automatically derived cardiac indices showed good agreement with manually derived ones, with small biases and narrow limits of agreement, particularly for EDV, ESV and LVEF. Moreover, the average analysis time was found to be nearly 150 times faster than manual contouring, which may facilitate its introduction in a clinical setting.

Acknowledgements

The authors acknowledge funding support from FCT - Fundação para a Ciência e Tecnologia, Portugal, in the scope of the PhD grant SFRH/BD/93443/2013, and the EU (FP7) framework program, for the financial support of DOPPLER-CIP project (grant number 223615).

References

1. Miller CA, Pearce K, Jordan P, Argyle R, Clark D, Stout M, et al. Comparison of real-time three-dimensional echocardiography with cardiovascular magnetic resonance for left ventricular volumetric assessment in unselected patients. European Heart Journal–Cardiovascular Imaging. 2012;13(2):187-95.
2. Miller CA, Jordan P, Borg A, Argyle R, Clark D, Pearce K, et al. Quantification of left ventricular indices from SSFP cine imaging: Impact of real-world variability in analysis methodology and utility of geometric modeling. Journal of Magnetic Resonance Imaging. 2012;37(5):1213-22.
3. Petitjean C, Dacher JN. A review of segmentation methods in short axis cardiac MR images. Medical image analysis. 2011;15(2):169-84.
4. Hautvast GL, Salton CJ, Chuang ML, Breeuwer M, O'Donnell CJ, Manning WJ. Accurate computer-aided quantification of left ventricular parameters: Experience in 1555 cardiac magnetic resonance studies from the Framingham Heart Study. Magnetic Resonance in Medicine. 2012;67(5):1478-86.
5. Anderson J, Weaver AN, Horne BD, Jones HU, Jelaco GK, Cha JA, et al. Normal cardiac magnetic resonance measurements and interobserver discrepancies in volumes and mass using the papillary muscle inclusion method. The Open General and Internal Medicine Journal. 2007;1:6-12.
6. Chuang ML, Gona P, Hautvast GL, Salton CJ, Blease SJ, Yeon SB, et al. Impact of left ventricular trabeculations and papillary muscles on measures of cavity volume and ejection fraction. Journal of Cardiovascular Magnetic Resonance. 2011;13(Suppl 1):P36.
7. van Geuns RJM, Baks T, Gronenschild EH, Aben J-PM, Wielopolski PA, Cademartiri F, et al. Automatic Quantitative Left Ventricular Analysis of Cine MR Images by Using Three-dimensional Information for Contour Detection. Radiology. 2006;240(1):215-21.
8. Alfakih K, Plein S, Thiele H, Jones T, Ridgway JP, Sivananthan MU. Normal human left and right ventricular dimensions for MRI as assessed by turbo gradient echo and steady-state free precession imaging sequences. Journal of Magnetic Resonance Imaging. 2003;17(3):323-9.
9. Luijnenburg SE, Robbers-Visser D, Moelker A, Vliegen HW, Mulder BJ, Helbing WA. Intra-observer and interobserver variability of biventricular function, volumes and mass in patients with congenital heart disease measured by CMR imaging. The International Journal of Cardiovascular Imaging. 2010;26(1):57-64.
10. Broussaud T, Nchimi A, editors. Estimation of left ventricle parameters from Cine MR data: Fully Automatic versus Manual Analysis. ISMRM 17th Annual Meeting; 2009; Hawaii.
11. Karamitsos TD, Hudsmith LE, Selvanayagam JB, Neubauer S, Francis JM. Operator Induced Variability in Left Ventricular Measurements with Cardiovascular Magnetic Resonance is Improved After Training. Journal of Cardiovascular Magnetic Resonance. 2007;9(5):777-83.
12. Queirós S, Barbosa D, Heyde B, Morais P, Vilaça JL, Friboulet D, et al. Fast automatic myocardial segmentation in 4D cine CMR datasets. Medical image analysis. 2014;18(7):1115-31.
13. Axel L, Sodickson DK. The Need for Speed: Accelerating CMR Imaging Assessment of Cardiac Function. JACC: Cardiovascular Imaging. 2014;7(9):893-5.
14. Codella NC, Cham MD, Wong R, Chu C, Min JK, Prince MR, et al. Rapid and accurate left ventricular chamber quantification using a novel CMR segmentation algorithm: a clinical validation study. Journal of Magnetic Resonance Imaging. 2010;31(4):845-53.
15. Souto M, Masip LR, Couto M, Suárez-Cuenca JJ, Martínez A, Tahoces PG, et al. Quantification of Right and Left Ventricular Function in Cardiac MR Imaging: Comparison of Semiautomatic and Manual Segmentation Algorithms. Diagnostics. 2013;3(2):271-82.
16. Lu Y-L, Connelly KA, Dick AJ, Wright GA, Radau PE. Automatic functional analysis of left ventricle in cardiac cine MRI. Quantitative imaging in medicine and surgery. 2013;3(4):200.

17. Tufvesson J, Ubachs JF, Steding-Ehrenborg K, Carlsson M, Arheden H, Heiberg E. Validation of an algorithm for left ventricular segmentation in 150 patients shows potential for further development towards fully automatic segmentation. Journal of Cardiovascular Magnetic Resonance. 2013;15(Suppl 1):E30.
18. Tufvesson J, Hedström E, Steding-Ehrenborg K, Carlsson M, Arheden H, Heiberg E. Validation and Development of a New Automatic Algorithm for Time-Resolved Segmentation of the Left Ventricle in Magnetic Resonance Imaging. BioMed Research International. 2015.
19. Rademakers F, Engvall J, Edvardsen T, Monaghan M, Sicari R, Nagel E, et al. Determining optimal noninvasive parameters for the prediction of left ventricular remodeling in chronic ischemic patients. Scandinavian Cardiovascular Journal. 2013;47(6):329-34.
20. Kramer CM, Barkhausen J, Flamm SD, Kim RJ, Nagel E. Standardized cardiovascular magnetic resonance (CMR) protocols 2013 update. J Cardiovasc Magn Reson. 2013;15:91.
21. Schulz-Menger J, Bluemke DA, Bremerich J, Flamm SD, Fogel MA, Friedrich MG, et al. Standardized image interpretation and post processing in cardiovascular magnetic resonance: Society for Cardiovascular Magnetic Resonance (SCMR) Board of Trustees Task Force on Standardized Post Processing. Journal of Cardiovascular Magnetic Resonance. 2013;15.
22. Papavassiliu T, Kühl HP, Schröder M, Süselbeck T, Bondarenko O, Böhm CK, et al. Effect of endocardial trabeculae on left ventricular measurements and measurement reproducibility at cardiovascular MR imaging1. Radiology. 2005;236(1):57-64.
23. Hudsmith LE, Petersen SE, Francis JM, Robson MD, Neubauer S. Normal human left and right ventricular and left atrial dimensions using steady state free precession magnetic resonance imaging. Journal of Cardiovascular Magnetic Resonance. 2005;7(5):775-82.
24. Heiberg E, Sjögren J, Ugander M, Carlsson M, Engblom H, Arheden H. Design and validation of Segment-freely available software for cardiovascular image analysis. BMC medical imaging. 2010;10(1):1.
25. Klinke V, Muzzarelli S, Lauriers N, Locca D, Vincenti G, Monney P, et al. Quality assessment of cardiovascular magnetic resonance in the setting of the European CMR registry: description and validation of standardized criteria. J Cardiovasc Magn Reson. 2013;15:55.
26. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. The statistician. 1983:307-17.
27. van der Geest RJ, Lelieveldt BP, Angelié E, Danilouchkine M, Swingen C, Sonka M, et al. Evaluation of a new method for automated detection of left ventricular boundaries in time series of magnetic resonance images using an active appearance motion model. Journal of Cardiovascular Magnetic Resonance. 2004;6(3):609-17.
28. Sigmund AM, Usman A, Wasielewski M, Akhavan AR, Giri S, Stratton MA, et al. Semi-automatic inline calculation of left ventricular function using cardiac MRI (CMR). Journal of Cardiovascular Magnetic Resonance. 2013;15(Suppl 1):P80.
29. Steen H, Nasir K, Flynn E, El-Shehaby I, Lai S, Katus HA, et al. Is magnetic resonance imaging the reference standard for cardiac functional assessment? Factors influencing measurement of left ventricular mass and volumes. Clinical Research in Cardiology. 2007;96(10):743-51.
30. Lang RM, Badano LP, Afilalo J, Mor-Avi V, Armstrong A, Ernande L, et al. Recommendations for cardiac chamber quantification by echocardiography: An update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. Eur Heart J Cardiovasc Imaging. 2014:in press.
31. Bonow RO, Ganiats TG, Beam CT, Blake K, Casey DE, Goodlin SJ, et al. ACCF/AHA/AMA-PCPI 2011 Performance Measures for Adults With Heart FailureA Report of the American College of Cardiology Foundation/American Heart Association Task Force on Performance Measures and the American Medical Association–Physician Consortium for Performance Improvement. Journal of the American College of Cardiology. 2012;59(20):1812-32.
32. Pibarot P, Larose É, Dumesnil J. Imaging of Valvular Heart Disease. Canadian Journal of Cardiology. 2013;29(3):337–49.

33. Coulden R, Sonnex E. Inter-observer variation in LV analysis in a dedicated CMR unit: the impact of audit and consensus guideline on reproducibility. Journal of Cardiovascular Magnetic Resonance. 2014;16(Suppl 1):P372.
34. Barbosa D, Heyde B, Cikes M, Dietenbeck T, Claus P, Friboulet D, et al. Real-time 3D interactive segmentation of echocardiographic data through user-based deformation of B-spline explicit active surfaces. Computerized Medical Imaging and Graphics. 2014;38(1):57-67.
35. Sugeng L, Mor-Avi V, Weinert L, Niel J, Ebner C, Steringer-Mascherbauer R, et al. Quantitative Assessment of Left Ventricular Size and Function Side-by-Side Comparison of Real-Time Three-Dimensional Echocardiography and Computed Tomography With Magnetic Resonance Reference. Circulation. 2006;114(7):654-61.
36. Trieb T, Glodny B, Scheiblhofer M, Wolf C, Metzler B, Pachinger O, et al. Inter- and intra-rater reproducibility of semiautomatic determination of volume parameters in cardiac magnetic resonance imaging. European journal of radiology. 2008;68(3):476-86.
37. Caudron J, Fares J, Lefebvre V, Vivier P-H, Petitjean C, Dacher J-N. Cardiac MRI assessment of right ventricular function in acquired heart disease: Factors of variability. Academic Radiology. 2012;19(8):991-1002.
38. Groth M, Muellerleile K, Klink T, Säring D, Halaj S, Folwarski G, et al. Improved Agreement between Experienced and Inexperienced Observers using a Standardized Evaluation Protocol for Cardiac Volumetry and Infarct Size Measurement. Fortschr Röntgenstr. 2012;184(12):1131-7.
39. Van der Geest RJ, Buller VG, Jansen E, Lamb HJ, Baur LH, van der Wall EE, et al. Comparison between manual and semiautomated analysis of left ventricular volume parameters from short-axis MR images. Journal of computer assisted tomography. 1997;21(5):756-65.
40. Hayes C, Guehring J, Jolly M, Schmidt M, Wintersperger B, Bhargava A, et al., editors. Fully automatic segmentation of the left ventricle in cardiac cine MR images. Proc International Society for Magnetic Resonance in Medicine; 2007.
41. Hegde VA, Mikolich JR, Doyle M, Rathi VK, Yamrozik JA, Williams RB, et al. Cardiac magnetic resonance imaging in today's economic climate; a cost effectiveness analysis. Journal of Cardiovascular Magnetic Resonance. 2011;13(Suppl 1):M12.

Text Tables

Table 1 – Patient characteristics.

	Number (%) $n = 318$
Age	64 ± 8.3
Gender (% male)	236 (74%)
Heart rate	66 ± 12.3
Risk factors:	
Hypertension	189 (59%)
Diabetes	44 (14%)
Smoking	27 (8%)
Family History	148 (47%)
Medical history	
Previous myocardial infarction	119 (37%)
Chronic renal disease	4 (1%)
Previous stroke	16 (5%)
Previous percutaneous coronary intervention	124 (39%)
Previous coronary artery bypass surgery	36 (11%)
Clinical heart failure - New York Heart Association (NYHA)	
Class I	161 (51%)
Class II	103 (32%)
Class III	54 (17%)
Class IV	0 (0%)

Table 2 – Typical scan parameters from the CMR image acquisition at the different sites.

Site	**LIO**	**RHS**	**KUL**	**KCL**
# patients	176	63	61	18
Scanner	Philips Achieva	Siemens Avanto	Philips Achieva	Philips Achieva
Field strength (T)	1.5	1.5	1.5	3.0
Echotime (ms)	1.77	1.25	1.33	1.32
Repetition time (ms)	3.55	20.5	2.70	2.65
Flip angle (degrees)	60	80	60	40
Acquired temporal resolution (ms)	39.05	34.08	83.52	55.65
Acquired spatial resolution (mm^2)	1.50 × 1.90	1.40 × 2.30	2.00 × 2.22	1.70 × 2.00
Slice thickness (mm)	8	6	8	8
Slice gap (mm)	2	4	0	0
Reconstructed resolution (mm^2)	1.20 × 1.20	1.40 × 1.40	1.25 × 1.25	1.15 × 1.15
Reconstructed timeframes	30	24	30	30
Slice per breath-hold	2	1	2	2

Note that the actual settings differed from patient to patient.

LIO – Linköping University; RHS – Oslo University Hospital; KUL – KU Leuven; KCL – King's College London.

Table 3 – 3D+time automatic segmentation performance (average perpendicular distance, APD) for endo and epicardial contours compared to manual tracings (# = 302 datasets). APD measures between manual tracings for intra- (# = 15 datasets) and interobserver (# = 32 datasets) variability analyses are also reported.

	APD_{endo} (mm)	**APD_{epi} (mm)**
Automatic vs Manual		
ED	1.52 ± 0.76	2.18 ± 0.91
ES	2.37 ± 0.90	3.42 ± 1.11
All	1.92 ± 0.71	2.76 ± 0.90
Intraobserver analysis		
ED	0.78 ± 0.23	0.85 ± 0.30
ES	0.96 ± 0.20	1.04 ± 0.27
All	0.86 ± 0.19	0.94 ± 0.25
Interobserver analysis		
ED	1.22 ± 0.45	1.25 ± 0.30
ES	1.70 ± 0.64	1.55 ± 0.45
All	1.44 ± 0.41	1.39 ± 0.31

Table 4 – Bland-Altman analysis for clinical cardiac indices between automatic segmentation and manual delineations (# = 302 datasets). Bias and limits of agreement for intra- (# = 15 datasets) and interobserver (# = 32 datasets) variability analyses are also reported.

	Automatic vs Manual		**Intraobserver analysis**		**Interobserver analysis**	
	Absolute	**Relative**	**Absolute**	**Relative**	**Absolute**	**Relative**
EDV	-4.08[a] ± 8.98	-2.61[a] ± 5.80	-1.13 ± 4.99[b]	-0.63 ± 3.66[b]	3.40 ± 9.52	2.74 ± 6.49
ESV	1.18[a] ± 9.74	1.97[a] ± 16.68	-1.41 ± 3.12[b]	-2.17 ± 5.49[b]	-4.81[a] ± 6.20[b]	-7.60[a] ± 7.72[b]
SV	-5.26[a] ± 9.08	-5.45[a] ± 11.13	0.27 ± 3.49[b]	0.85 ± 5.19[b]	8.21[a] ± 7.71	11.34[a] ± 10.11
LVM	20.41[a] ± 20.38	21.33[a] ± 21.14	-0.90 ± 7.59[b]	-1.59 ± 8.53[b]	-8.84[a] ± 16.05	-8.87[a] ± 14.85
EF	-1.53[a] ± 4.93	-2.85[a] ± 10.41	0.66 ± 1.42[b]	1.50 ± 3.78[b]	4.21[a] ± 2.65[b]	8.28[a] ± 5.71[b]

[a] $p < 0.05$, two-tailed paired t-test against zero.
[b] $p < 0.05$, two-tailed F-test against "Automatic vs Manual" analysis.

Table 5 – Comparison of clinical parameters computation performance against literature (# - number of datasets).

	Technique	**#**	**EDV (mL)**	**ESV (mL)**	**SV (mL)**	**LVM (g)**	**LVEF (%)**
Proposed	Automatic	302	-4.08 ± 8.98	1.18 ± 9.74	-5.26 ± 9.08	20.41 ± 20.38	-1.53 ± 4.93
van der Geest *et al.* [39]	Semi-automated	20	-5.5 ± 9.7	-3.6 ± 6.5	-	7.3 ± 20.6	1.7 ± 4.1
van der Geest *et al.* [27]	Automatic	17	-2.9 ± 13.2	-5.1 ± 18.9	-	-1.2 ± 14.1	0.1 ± 6.7
van Geuns *et al.* [7]	Semi-automated	25	-8.15 ± 11.46	-5.95 ± 6.34	-	7.19 ± 15.00	1.6 ± 3.5
Hayes *et al.* [40]	Automatic	20	-8.29 ± 10.38	-5.05 ± 8.32	-	-	0.76 ± 3.91
Brossaud *et al.* [10]	Fully automatic	130	-11.9 ± 10.2	-8.4 ± 6.9	-3.3 ± 8.3	-	3.7 ± 4.5
Codella *et al.* [14]	Automatic[†]	151	4.0 ± 6.8	1.4 ± 5.5	2.6 ± 5.3	-	0.6 ± 2.3
Souto *et al.* [15]	Semi-automated	52	-4.1 ± 19.0	-3.7 ± 13.5	-	-	1.1 ± 7.0
Lu *et al.* [16]	Automatic	133	-1.69 ± 12.76	-1.51 ± 11.30	-	-0.66 ± 14.72	0.02 ± 5.93
Tufvesson *et al.* [17]	Automatic[*]	150	-14.4 ± 9.0	-10.8 ± 8.7	-	11.3 ± 14.4	2.5 ± 2.7
Tufvesson *et al.* [18]	Automatic	49	-11 ± 11	1 ± 10	-12 ± 8	4 ± 15	-3 ± 4

[†] Manual definition of the valve annulus in non-circumferential myocardium and manual corrections allowed.
[*] Most basal slice and basal slices with a non-circumferential myocardium were manually contoured.

Figure Legends

Figure 1 – Dataflow for automated LV segmentation algorithm.

Figure 2 – Automatic segmentation results at the ED phase for 5 representative cases covering the range between the best (A) and the worst (E) segmentation results (red: endocardium; green: epicardium; yellow: ground truth).

Figure 3 - Automatic segmentation results at the ES phase for the cases illustrated in Figure 2 (red: endocardium; green: epicardium; yellow: ground truth).

Figure 4 – Bland-Altman analysis for (A) end-diastolic volume (EDV); (B) end-systolic volume (ESV); (C) LV mass (LVM); and (D) ejection fraction (EF). Both bias (mean difference) and 95% limits of agreement (± 1.96 SD around the mean difference) are depicted.

Figure 5 – Bland-Altman analysis of (A-D) intra- and (E-H) interobserver variability analyses for (A,E) end-diastolic volume (EDV); (B,F) end-systolic volume (ESV); (C,G) LV mass (LVM); and (D,H) ejection fraction (EF). Both bias (mean difference) and 95% limits of agreement (± 1.96 SD around the mean difference) are depicted.